Analysis of Sea Clutter Data

Irina Antipov

DSTO-TR-0647

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

Analysis of Sea Clutter Data

Irina Antipov

Tactical Surveillance System Division
Electronic and Surveillance Research Laboratory

DSTO-TR-0647

ABSTRACT

This report presents the results of comparative analysis of the models which have been applied to sea clutter amplitude distribution, and the methods for estimation of their parameters. More detail consideration is given to the K-distribution as the most appropriate model for sea clutter in the low Probability of False Alarm (PFA) region.

Published by

DSTO Electronic and Surveillance Research Laboratory
PO Box 1500
Salisbury South Australia 5108

Telephone: (08) 8259 5200
Fax: (08) 8259 5200

AR No. 010-489
March 1998

APPROVED FOR PUBLIC RELEASE

Analysis of Sea Clutter Data

Executive Summary

This report represents the results of a study performed under the task ADA95/080-TSSD support for Project AIR 5276. Under this task, TSSD is providing advice to the RAAF on several aspects of the upgrade of the P3C Orions with a focus on the Israeli designed EL/M 2022 maritime surveillance radar system being installed as part of this upgrade. The performance of maritime surveillance radars is adversely affected by returns from the sea surface - sea clutter, therefore understanding of the behaviour and properties of sea clutter is important for validation and optimisation of EL/M 2022 detection performance prediction in support of AIR 5276.

For radars in which the resolution cell dimensions are much greater than the sea swell wavelength and for grazing angles greater than about 10^o, it is well known that the clutter amplitude is Rayleigh distributed. As the radar resolution is increased and/or for smaller grazing angles, the clutter amplitude distribution is observed to develop a longer "tail" and displays a larger standard deviation-to-mean ratio than would be predicted under the Rayleigh distribution.

The report presents the results of comparative analysis of the main non-Rayleigh models (Log-Normal, Weibull and K-distribution) which have been applied to sea clutter amplitude distribution, and the methods for estimation of their parameters.

The methods are compared against a single data set obtained from the data base of recording signals of the cliff-top positioned INGARA system. It has been shown that the K-distribution is the most appropriate model for sea clutter in the low Probability of False Alarm (PFA) region.

Several existing methods for estimation of the parameters of the K-distribution have been analysed, and some recommendations about their implementation have been suggested.

Except for the K-distribution, non-Rayleigh distributions are not derived from a physical model or clutter scattering mechanism. Their choice and validation are based only on their agreement with experimental data. Although the amplitude distribution is sufficient to predict the performance of fixed threshold single pulse detection, other forms of processing require knowledge of the correlation properties. The effect of describing sea clutter by using the compound form of the K-distribution is to allow a greater variety of processing schemes to be accurately evaluated; these include pulse to pulse integration and cell averaging Constant False Alarm Rate (CFAR) processing.

Authors

Irina Antipov
Tactical Surveillance System Division

Dr. Irina Antipov received her degree in 1988 from St-Petersburg Academy of Aerospace Instrumentation, Russia. From 1996 she has worked in Tactical Surveillance System Division as a Research Scientist. Her research interests are in the modelling and simulation of sea clutter for a variety of purposes including: investigating the performance of maritime radar systems, adapting radar systems to operate in clutter environment, investigating concepts of operation of maritime radar systems.

Contents

List of figures

List of tables

1. Introduction

The performance of maritime surveillance radars is adversely affected by returns from the sea surface - sea clutter. Although there exists much theoretical and experimental work in the literature [1-22], our understanding of sea clutter behaviour is still incomplete. Thus, different authors use different models for the sea clutter amplitude distribution, and suggest different methods for estimating the parameters of these distributions. This report presents the results of comparative analysis of the models which have been applied to sea clutter amplitude distribution, and the methods for estimation of their parameters. More detailed consideration is given to the K-distribution as the most appropriate model for sea clutter in low Probability of False Alarm (PFA) region.

With the exseption of K-distribution, non-Rayleigh distributions are not derived from a physical model or clutter scattering mechanism. Their choice and validation are based only on their agreement with experimental data. Although the amplitude distribution is sufficient to predict the performance of fixed threshold single pulse detection, other forms of processing require knowledge of the correlation properties. The effect of describing sea clutter by using the compound form of the K-distribution [7-9] is to allow a greater variety of processing schemes to be accurately evaluated; these include pulse to pulse integration and cell averaging Constant False Alarm Rate (CFAR) processing. It has been found [7] that generally the correlation properties, as expressed in the compound form of the K-distribution, have much more impact on processing performance than the non-Rayleigh amplitude distribution alone.

2. Analysis of sea clutter data

For radars in which the resolution cell dimensions are much greater than the sea swell wavelength, and for grazing angles greater than about 10^o, it is well known that the clutter amplitude is Rayleigh distributed. This is a consequence of the Central Limit Theorem since the returns can be thought of as being the vector sum of randomly phased components from a large number of independent scatterers. The clutter returns have a fairly short temporal decorrelation period (typically of the order of 10 ms) and may be fully decorrelated from pulse from pulse by the use of frequency agility with the steps at least the transmitted pulse bandwidth. There is no inherent correlation of returns in range, the sole determinant of this being the transmitted pulse length [1,2].

As the radar resolution is increased and/or for smaller grazing angles, the clutter amplitude distribution is observed to develop a longer "tail" (higher number of large amplitude values) and displays a larger standard-deviation-to-mean ratio than would be predicted under the Rayleigh distribution. The returns are often described as becoming "spiky". The temporal and spatial correlation characteristics of the clutter also change. Clutter spikes having temporal decorrelation periods of several seconds

or more are often observed, and the clutter returns are no longer fully decorrelated from pulse to pulse by the use of frequency agility. Spatial correlation between echo amplitudes from adjacent resolution cells appears, which is generally more pronounced for cell sizes smaller than the scale of large features on the sea surface (such as the wavelength of the dominant ocean gravity wave).

Partially successful models for the amplitude distributions that have been applied to such data sets include the Log-Normal and Weibull. These models have resulted from empirical studies of the sea clutter and are not based on any physical understanding of the clutter returns. Another distribution that has become increasingly popular is the compound K-distribution, which has the advantage over other models in that it does have a theoretical justification and physical interpretation. The following sections briefly describe the distributions, their capability to represent the real sea clutter statistical properties, and the methods for estimating of their parameters.

Throughout this report, it will be assumed that independent samples are available for parameter estimation. As high resolution radar images contain correlations resulting from underlying cross-section modulations, this assumption necessitates subsampling of the data to remove any correlations before parameter estimation is performed. Therefore, initially detailed analysis of sea clutter correlation properties for high resolution radars has to be done, in order to choose the correct way of subsampling the data.

It has been found from practical measurements over a wide range of conditions that while microwave signals are primarily scattered by capillary waves of the ocean, the undulating structure of the ocean gravity waves causes variations in the mean power level scattered from a given patch [3-10]. Therefore, the clutter returns for high resolution radars can be well modelled by two components.

The first component models scattering effects due to small scale structure (such as capillary wave or ripples) and has a Rayleigh distribution. Temporal and spatial correlation properties of this component are :

- the Rayleigh component decorrelates over a few milliseconds (for fixed frequency operation this component of clutter returns will typically decorrelate over periods of 5 to 10 ms) and can be decorrelated from pulse to pulse by frequency agility with steps greater than or equal to the pulse bandwidth;

- the small scale features at two spatially separated patches are uncorrelated.

The second component models the local mean of the first component, ie it models scattering effects due to large scale structures (such as gravity waves). The mean level varies in time, and from cell to cell, having a power which is Gamma distributed. Temporal and spatial correlation properties of this component are different from those of the Rayleigh component:

- the Gamma component is unaffected by frequency agility and varies only slowly with time. As a result, independent samples of the underlying mean level are not achieved during the dwell time of conventional maritime surveillance radars. Moreover, the returns from a particular spike may be highly correlated even over several scans [7];

- spatial correlation of scattered power cannot be ignored if the physical separation of the patches is less than the correlation length of large scale structures.

Statistically independent samples of the sea clutter amplitude distribution may be obtained by:

- taking from a resolution cell, echo samples that are separated in time by intervals greater than the decorrelation time of the Gamma component. If the time separation of two echo samples is greater than the decorrelation time of the Rayleigh component but less than the decorrelation time of the Gamma component, the Rayleigh amplitude components for different returns from the cell will be independent, but the mean values will be correlated ;

- taking echo samples from resolution cells that are separated by distances greater than the decorrelation length of the large scale effects. If the spatial separation of two nonoverlapping resolution cells is less than the large scale decorrelation length, the individual Rayleigh amplitude components from the cells will be independent, but the underlying mean values will be correlated.

Ideally, parameter estimation would take into account the correlations in the data, and would use all the available samples. However, the analysis of such correlated variables is difficult and beyond the scope of this report.

2.1 Log-Normal Distribution

Sea clutter statistics approach those of the Log-Normal distribution when a high resolution and horizontal polarisation radar sees sea clutter at low grazing angles ($\varphi < 5^{o}$) [3,11,12].

In general, the Log-Normal model tends to overestimate the dynamic range of the actual clutter distribution while the Rayleigh model tends to underestimate the dynamic range.

From the detection standpoint, it can be said that the Log-Normal distribution represents a worst case distribution compared with the Rayleigh distribution which represents the best case.

2.1.1 Statistical characteristics of the Log-Normal distribution

The Probability Density Function (PDF) of the Log-Normal distribution is given by

$$f_L(a) = \frac{1}{\sqrt{2\pi}\sigma a}\exp\{-\frac{(\ln(a)-\mu)^2}{2\sigma^2}\}, \qquad a>0. \tag{1}$$

A random variable a, is Log-Normal if and only if $\ln(a)$ is Normally distributed with mean μ and variance σ^2; in our case, a signifies the random amplitude returns of the sea clutter.

The model Cumulative Distribution Function (CDF) then may be expressed in terms of the Normal distribution, that is:

$$F_L(a_0) = P(a \le a_0) = P(\ln(a) \le \ln(a_0)) = \Phi\{[\ln(a_0)-\mu]/\sigma\}, \qquad a>0, \tag{2}$$

where $\Phi(z) = \int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}\exp\{-\frac{t^2}{2}\}dt.$

The false alarm rate for a given threshold a_0 is given by expression

$$P(a > a_0) = 1 - F_L(a_0). \tag{3}$$

The r^{th} moment of a can be shown to be

$$\mu_r = E(a^r) = \exp(r\mu + \frac{1}{2}r^2\sigma^2). \tag{4}$$

Therefore, the mean and variance for the Log-Normal distribution are:

$$E(a) = \exp(\mu + \frac{1}{2}\sigma^2), \tag{5}$$

$$Var(a) = \exp(2\mu+\sigma^2)\{\exp(\sigma^2)-1\}. \tag{6}$$

2.1.2 Estimation of the parameters of the Log-Normal distribution

A standard approach to parameter estimation is to use the maximum likelihood (ML) solution, which provides optimum parameter estimates in the sense that these estimates are the most probable parameter values, given the data (i.e. the sample values) but no prior knowledge [13-15].

If n independent samples $a_1,a_2,...,a_n$ are drawn from a distribution with m parameters $\theta_1,\theta_2,...,\theta_m$, then the joint PDF of $a_1,a_2,...,a_n$ is the product of the marginal PDFs:

$$f(a_1,a_2,...,a_n;\theta_1,\theta_2,...,\theta_m) = f(a_1;\theta_1,\theta_2,...,\theta_m)f(a_2;\theta_1,\theta_2,...,\theta_m)...f(a_n;\theta_1,\theta_2,...,\theta_m)$$

This PDF when expressed as a function of $\theta_1,\theta_2,...,\theta_m$ is called the likelihood function:

$$L_n(\theta_1,\theta_2,...,\theta_m;a_1,a_2,...,a_n) = \prod_{i=1}^{n} f(a_i;\theta_1,\theta_2,...,\theta_m). \quad (7)$$

The ML estimate of $\theta_1,\theta_2,...,\theta_m$ is the set of values $\hat{\theta}_1,\hat{\theta}_2,...,\hat{\theta}_m$ that maximises the value of the likelihood function. The ML estimate is the "most likely" set of parameter values given the observed data.

Since the logarithm function is a monotonic function, the ML estimate will also maximise the log-likelihood function:

$$\ln[L_n(\theta_1,\theta_2,...,\theta_m;a_1,a_2,...,a_n)] = \sum_{i=1}^{n} \ln[L_n(\theta_1,\theta_2,...,\theta_m;a_i)]. \quad (8)$$

In many cases it is easier to obtain a closed-form solution with the log-likelihood function. If numerical maximisation must be used, elimination of the product gives more stable results.

The log-likelihood function for the Log-Normal distribution is given by

$$\ln[L_n(\mu,\sigma^2;a_1,a_2,...,a_n)] = \sum_{i=1}^{n} \ln[L_n(\mu,\sigma^2;a_i)]. \quad (9)$$

Taking the derivative of (9) with respect to the parameters μ and σ^2 and setting each to zero gives the following system of equations for the ML estimates of the parameters, which are denoted by $\hat{\mu}_{ML}$ and $\hat{\sigma}^2_{ML}$ respectively:

$$\begin{cases} \dfrac{\partial}{\partial\mu}[\sum_{i=1}^{n} \ln(L_n(\mu,\sigma^2;a_i))]|_{\mu=\hat{\mu}_{ML}} = 0 \\ \dfrac{\partial}{\partial\sigma^2}[\sum_{i=1}^{n} \ln(L_n(\mu,\sigma^2;a_i))]|_{\sigma^2=\hat{\sigma}^2_{ML}} = 0 \end{cases}. \quad (10)$$

ML estimates have the following asymptotic properties [13]:

1) ML estimates are consistent, i.e. they converge in probability to the true value as the sample size increases to infinity;

2) ML estimates are asymptotically efficient, ie. they approach the Cramer-Rao bound as the sample size or the signal-to-noise ratio increases to infinity.

The ML estimates of parameters of the Log-Normal distribution μ and σ^2 are:

$$\hat{\mu}_{ML} = \frac{1}{n}\sum_{i=1}^{n} \ln(a_i), \tag{11}$$

$$\hat{\sigma}^2_{ML} = \frac{1}{n}\sum_{i=1}^{n} [\ln(a_i) - \hat{\mu}_{ML}]^2. \tag{12}$$

Formula (12) gives a consistent but biased estimate of σ^2. An unbiased estimate of σ^2 is defined by

$$(\hat{\sigma}^2_{ML})_{unb} = \frac{1}{(n-1)}\sum_{i=1}^{n} [\ln(a_i) - \hat{\mu}_{ML}]^2. \tag{13}$$

2.2 Weibull Distribution

The Weibull PDF is a two-parameter distribution, of which the Rayleigh distribution is a special case. The first parameter of the distribution, a shape parameter, relates to the skewness to the distribution; whereas the second parameter, a scale parameter, scales the distribution, as the name implies. This distribution is mathematically convenient as it allows the skewness of the distribution to be changed with a single parameter to match the characteristics of the data.

The Weibull PDF is known to represent sea clutter quite well at low grazing angles and/or at high-resolution situations: the choice of appropriate values of the shape and scale parameters allows the simulation of returned echo signals with required spikiness and power characteristics.

The Weibull clutter model [15-19] offers the potential to accurately represent the real clutter distribution over a much wider range of conditions than either the Log-Normal or Rayleigh model. By appropriately adjusting its parameters, the Weibull distribution can be made to approach either the Rayleigh (a member of the Weibull family) or Log-Normal distribution. From the detection standpoint, it can be said that the Weibull distribution represents an intermediate model that may more accurately represent the real detection performance in clutter than either Rayleigh or Log-Normal distribution.

2.2.1 Statistical characteristics of the Weibull distribution

In the Weibull clutter model, the amplitude PDF is given by

$$f_W(a) = \frac{\gamma}{\varpi}(\frac{a}{\varpi})^{\gamma-1}\exp\{-(\frac{a}{\varpi})^{\gamma}\}, \qquad a \geq 0, \varpi \geq 0, \gamma \geq 0, \tag{14}$$

where γ is the shape parameter and ϖ is the scale parameter of the Weibull distribution, respectively.

The Weibull distribution reduces to the Rayleigh distribution for $\gamma = 2$ (which has been a common model for low-resolution sea clutter [3]). Smaller values of γ increase the skewness of this distribution, and allow the simulation of spiky clutter.

The CDF of the Weibull distribution is given by

$$F_W(a) = 1 - \exp\{-(\frac{a}{\varpi})^{\gamma}\}. \tag{15}$$

The r^{th} moment of a is

$$\mu_r = E(a^r) = \varpi^r \Gamma(1+\frac{r}{\gamma}), \tag{16}$$

where $\Gamma(z)$ is the Gamma function.

Therefore, the mean and variance for the Weibull distribution are:

$$E(a) = \varpi\Gamma(1+\frac{1}{\gamma}), \tag{17}$$

$$Var(a) = \varpi^2[\Gamma(1+\frac{2}{\gamma}) - \Gamma^2(1+\frac{1}{\gamma})]. \tag{18}$$

2.2.2 Estimation of the parameters of the Weibull distribution

The maximum likelihood equations for the parameters of the Weibull distributions are

$$\begin{cases} \dfrac{\partial}{\partial\mu}[\sum_{i=1}^{n}\ln(L_n(\mu,\sigma^2;a_i))]|_{\mu=\hat{\mu}_{ML}} = 0 \\ \dfrac{\partial}{\partial\sigma^2}[\sum_{i=1}^{n}\ln(L_n(\mu,\sigma^2;a_i))]|_{\sigma^2=\hat{\sigma}^2_{ML}} = 0 \end{cases}. \tag{19}$$

After elementary transformations, the above equations may be put in the form

$$\begin{cases} \dfrac{\sum_{i=1}^{n} a_i^{\hat{\gamma}_{ML}} \ln(a_i)}{\sum_{i=1}^{n} a_i^{\hat{\gamma}_{ML}}} - \dfrac{1}{\hat{\gamma}_{ML}} = \dfrac{1}{n}\sum_{i=1}^{n} \ln(a_i) \\ \hat{\varpi}_{ML} = \left[\dfrac{\sum_{i=1}^{n} a_i^{\hat{\gamma}_{ML}}}{n} \right]^{\frac{1}{\hat{\gamma}_{ML}}} \end{cases}, \tag{20}$$

which are not of closed form. An iterative procedure such as the Newton-Raphson technique may be utilised to solve these equations, yielding the ML estimates of the shape and scale parameters of the Weibull distribution. Convergence can be somewhat slow, and simpler closed form solutions, using either moments or order statistics [18], are available. Unfortunately, both simpler techniques exhibit a bigger variance of the estimated parameters compared with the ML method.

Menon proposed an estimation procedure which has the attractive feature of leading to a chosen false alarm rate for all values of γ and ϖ [19]. The estimators are

$$\hat{\gamma}_{Men} = \left\{ \frac{6}{\pi^2} \frac{n}{n-1} \left[\frac{1}{n}\sum_{j=1}^{n} \left(\ln(a_i)\right)^2 - \left(\frac{1}{n}\sum_{j=1}^{n} \ln(a_i) \right)^2 \right] \right\}^{-1/2}, \tag{21}$$

$$\hat{\varpi}_{Men} = \exp\left[\frac{1}{n}\sum_{j=1}^{n} \ln(a_i) + 0.5772\hat{\gamma}_{Men}^{-1} \right]. \tag{22}$$

2.3 K-distribution

The statistical results of many experiments in recent years provide evidence that the K-distribution can serve as a limiting distribution for sea clutter [3,4,7-12].

The K-distribution is based on an underlying physical model that treats the received signal as a superposition of returns from a number of independent patches or scatterers, illuminated by the radar beam [5]. The effective number of scatterers along with their relative bunching is critical in determining the overall statistics of the received data.

Analysis of the sea clutter data displays two dominant components with differing correlation times, contributing to the amplitude distribution [5,8-10,20,21]. The fast

varying component, which can be identified with the changing interference between capillary waves, has a correlation time in the order of milliseconds and can be decorrelated by the use of frequency agility. The slow varying component which can be associated with the gross wave structure of the sea surface, has a correlation time of the order of seconds and is unaffected by frequency agility. The spatial correlation of the components is also different. The interference or speckle component has range correlation commensurate with the pulse length and is in all ways similar to noiselike clutter, as its amplitude distribution confirms. The second component has considerable spatial correlation, depending upon aspect, which displays periodic effects and is coupled to the temporal correlation.

2.3.1 Statistical characteristics of the K-distribution

In the K-distribution model, the overall amplitude of the sea clutter return is represented as the product of two independent random variables:

$$a = y\upsilon, \tag{23}$$

where y is the voltage envelope modulation process, which has a long correlation time and spatial and temporal structure, and υ is the speckle voltage, which can be decorrelated by frequency agility.

The results from the averaged clutter returns show that $f(y)$, the PDF of y, fits well to the generalised Chi-distribution over a wide range of radar parameters and sea conditions:

$$f(y) = \frac{2d^{2\nu} y^{2\nu-1}}{\Gamma(\nu)} \exp(-d^2 y^2), \tag{24}$$

where $\Gamma(\nu)$ is the Gamma function, ν is a shape parameter and d is a scale parameter such that $d^2 = \frac{\nu}{E(y^2)}$ where $E(y^2)$ is the average power of the clutter. The value of ν depends on range, grazing angle, aspect angle, sea conditions and radar parameters.

The speckle component is well modelled by the Rayleigh distribution and has a mean level determined by the first slowly varying component:

$$f(a|y) = \frac{a\pi}{2y^2} \exp(-\frac{a^2\pi}{4y^2}). \tag{25}$$

The overall amplitude of the clutter is

$$f_K(a) = \int_0^\infty f(a|y) f(y) dy = \frac{2c}{\Gamma(\nu)} \left(\frac{ca}{2}\right)^\nu K_{\nu-1}(ca), \tag{26}$$

where $K_{\nu-1}(z)$ is an ν^{th}-order modified Bessel function of the second kind, hence the name k-distribution and $c = \sqrt{\pi d}$ is a scale parameter.

The CDF of the K-distribution is

$$F_K(a) = 1 - \frac{2}{\Gamma(\nu)} \left(\frac{ca}{2}\right)^\nu K_\nu(ca). \tag{27}$$

The r^{th} moment of a is

$$\mu_r = E(a^r) = \frac{2^r \Gamma(0.5r+1)\Gamma(0.5r+\nu)}{\Gamma(\nu) c^r}. \tag{28}$$

Therefore, the mean and variance for the K-distribution are

$$E(a) = \frac{2\Gamma(1.5)}{c} \frac{\Gamma(0.5+\nu)}{\Gamma(\nu)} = \frac{\sqrt{\pi}}{c} \frac{\Gamma(0.5+\nu)}{\Gamma(\nu)}, \tag{29}$$

$$Var(a) = \frac{4}{c^2} \frac{\Gamma(2)\Gamma(1+\nu)}{\Gamma(\nu)} - E^2(a) = \frac{4\nu}{c^2} - E^2(a). \tag{30}$$

Comparison with the other models shows that

- when the shape parameter of the K-distribution is equal to infinity, the K-distribution reduces to the Rayleigh distribution;
- the Log-Normal distribution is always spikier than the K-distribution;
- when the shape parameter is equal to 0.5, the K-distribution and the Weibull are identical. Over a large range of values of the shape parameter, they are very similar, with the K-distribution being slightly more spiky than the Weibull for larger values, and slightly less spiky for smaller values.

2.3.2 Estimation of the parameters of the K-distribution

Estimating the parameters of the K-distribution from a set of sample data to establish the threshold level has been found to be a non trivial problem. In practice, the value for the shape parameter ν of this distribution varies between the values of 0.1, corresponding to very spiky data and 20, corresponding to approximately Rayleigh

distributed data. The form of the expression for the K-distribution is difficult to work with and unfortunately the ML solution is analytically intractable, which means that either the solution must be obtained numerically, or an alternative estimation scheme must be used. A numerical solution has the advantage of providing the ML parameter estimates but is liable to be computationally inefficient in comparison with alternative schemes. On the other hand, alternative estimation schemes may be computationally efficient, but the parameter estimates that they provide may be subject to large errors. It is thus of interest to compare the parameter estimation errors of these alternative schemes with those of the ML solution, to facilitate assessment of the trade-off between computational efficiency and estimation accuracy. The following sections of the report describe existing parameter estimation methods for the K-distribution and compare their characteristics.

2.3.2.1 Maximum likelihood method

The log-likelihood function of a sample $\mathbf{a} = (a_1, a_2, \ldots, a_n)'$ from the K-distribution is given by the expression

$$\ln[L_n(\nu, c; \mathbf{a})] = n(1-\nu)\ln(2) + n(1+\nu)\ln(c) - n\ln\Gamma(\nu) + \nu\sum_{i=1}^{n}\ln(a_i) + \sum_{i=1}^{n}\ln[K_{\nu-1}(ca_i)].$$

The maximum likelihood equations for the parameters of the K-distributions are

$$\begin{cases} \dfrac{\partial}{\partial \nu}\left[\sum_{i=1}^{n}\ln(L_n(\nu, c; a_i))\right]\Big|_{\nu=\hat{\nu}_{ML}} = 0 \\ \dfrac{\partial}{\partial c}\left[\sum_{i=1}^{n}\ln(L_n(\nu, c; a_i))\right]\Big|_{c=\hat{c}_{ML}} = 0 \end{cases} . \tag{31}$$

As it is not possible to find analytical expressions for $\hat{\nu}_{ML}$ and $\hat{c}_{ML}$, the system (31) must be solved numerically. The ML estimates are obtained by using a numerical two-dimensional search in parameter space to locate the global maximum of the log-likelihood function. This procedure is tedious and the computational complexity increases as the number of samples is increased. The major factor in the overall computation time is the length of time it takes to evaluate the log-likelihood function [14]. Most of this time is spent computing the Bessel function for n independent different values. Even for small sample sizes ($n \leq 100$) it is difficult to implement the ML method from a computational point of view, and the amount of computations required to locate the maximum of the log-likelihood function makes this method impractical.

Nevertheless, the ML method provides a means of estimating the variance in the estimated parameters. In [14] it is shown that for the K-distribution the variance of the ML estimate $\hat{\nu}_{ML}$ satisfies

$$\hat{\sigma}^2_{\nu_{ML}} = \frac{\left.\frac{\partial^2 \Lambda(\nu,c;\mathbf{a})}{\partial c^2}\right|_{\nu=\hat{\nu}_{ML},c=\hat{c}_{ML}}}{\left[\left.\frac{\partial^2 \Lambda(\nu,c;\mathbf{a})}{\partial \nu \partial c}\right|_{\nu=\hat{\nu}_{ML},c=\hat{c}_{ML}}\right]^2 - \left[\left.\frac{\partial^2 \Lambda(\nu,c;\mathbf{a})}{\partial \nu^2}\right|_{\nu=\hat{\nu}_{ML},c=\hat{c}_{ML}}\right]\left[\left.\frac{\partial^2 \Lambda(\nu,c;\mathbf{a})}{\partial c^2}\right|_{\nu=\hat{\nu}_{ML},c=\hat{c}_{ML}}\right]}$$

where $\Lambda(\nu,c;\mathbf{a}) = \ln L_n(\nu,c;\mathbf{a})$.

Interchanging ν and c yields a similar relationship for $\hat{\sigma}^2_{\hat{c}_{ML}}$. There are no closed-form solutions for these relationships and, as a result, they must be evaluated numerically to obtain the estimated variances.

As the ML estimates can only be calculated by cumbersome numerical techniques, it is desirable to find an alternative estimation scheme that is easy to implement while providing near optimum error performance. Several alternative estimation schemes will be considered, which are based either on the matching of two different moments of the data, or on using the arithmetic and geometric sample means.

2.3.2.2 Methods of moments

A simple method of estimating the shape and scale parameters of the K-distribution using higher order moments is based on the result that moments of K-distributed random variable a are given by (28) and so any two sample moments of the data may be used for estimating these parameters. Thus, the ratio

$$\alpha_k = \frac{E(a^{2k})}{E^2(a^k)}, \qquad k=1,2,... \tag{32}$$

is independent of the scale parameter c and may be used for estimating the shape parameter ν. Then the estimate $\hat{\nu}$ and any one of the sample moments may be used for estimating the scale parameter c.

Given n independent clutter samples, the sample moments are obtained as

$$m_r = \frac{1}{n}\sum_{i=1}^{n} a_i^r, \tag{33}$$

where m_r is the r^{th} sample moment, n is the number of samples and a_i is the amplitude of r^{th} radar return.

Because the estimation of the shape and scale parameters of the K-distribution may be accomplished by equating theoretical expressions and calculated values for any two moments of the distribution and then solving for the two parameters, the question arises as to which two moments should be chosen. The choice will be governed by the errors in the resulting estimates, which can be quantified by calculating the variances, and hence the standard deviations, of the estimated values.

The methods of moments work well when the number of independent samples from the amplitude distribution is large (i.e. greater than about 1000), but the variance in estimates is large for smaller sample sizes.

2.3.2.2.1 Second and fourth sample moments based (Watts's) method

Results in which the parameters are estimated using the second and fourth sample moments are given by Watts [9]. Once these moments are known, then the following expressions can be used to calculate the values of the shape and scale parameters of the K-distribution with identical second and fourth moments as the sample - i.e. the shape and scale parameters for the K-distribution that most closely matches the distribution of the test sample:

$$\nu = \left(\frac{m_4}{2m_2^2} - 1 \right)^{-1}, \tag{34}$$

$$c = 2\sqrt{\frac{\nu}{m_2}}. \tag{35}$$

In the absence of thermal noise, when the clutter sample size is big enough, it has been found that the parameters ν and c obtained from using (34) and (35) are in good agreement with the true parameters of the K-distribution for observed sea clutter. However, these expressions do not provide high estimation accuracy of the K-distribution parameters when the level of thermal noise in the test sample is large. Thus, if a high level of noise is expected, the modified expressions for the estimation of the parameters ν and c have to be used.

2.3.2.2.2 Modified Watts's method

In the presence of high noise, the speckle component of the return is effectively modified by an increase in its average power. The new PDF of the speckle component is given by [10]

$$f_K(a|y) = \frac{a}{\sigma^2 + \frac{2y^2}{\pi}} \exp\left(-\frac{a^2}{2\sigma^2 + \frac{4y^2}{\pi}} \right) , \quad (36)$$

where a is the overall clutter-pulse-noise return and $2\sigma^2$ is the noise power level.

The CDF of this combined clutter and noise distribution is

$$F_{K+N}(a_0) = P(a > a_0) = \int_0^\infty \exp\left(\frac{-a_0}{2\sigma^2 + (4y^2/\pi)} \right) \frac{2d^{2\nu} y^{2\nu-1}}{\Gamma(\nu)} \exp(-d^2 y^2) dy \ . \quad (37)$$

The n^{th} moment of the distribution of clutter plus noise is given by

$$m_n = \int_0^\infty \left(2\sigma^2 + \frac{4y^2}{\pi} \right)^{n/2} \Gamma\left(\frac{n+2}{2} \right) \frac{2d^{2\nu}}{\Gamma(\nu)} y^{2\nu-1} \exp(-d^2 y^2) dy \ . \quad (38)$$

As a result, if the recorded data has a low clutter-to-noise ratio (CNR < 10 dB), then the resulting amplitude distribution will be significantly altered from a standard K-distribution with the added noise having the most effect on the low amplitude values of the distribution.

According to [10], if the recorded data is of high quality and a sufficiently large number of independent samples is available to estimate the higher moments accurately, it is possible to estimate both parameters of the K-distribution and CNR using the second, fourth and sixth moments of the recorded data:

$$\nu = \frac{18(m_4 - 2m_2^2)^3}{(12m_2^3 - 9m_2 m_4 + m_6)^2} \ , \quad (39)$$

$$2\sigma^2 = m_2 - [\frac{\nu}{2}(m_4 - 2m_2^2)]^{1/2} \ , \quad (40)$$

$$c = \sqrt{\frac{4\nu}{m_2 - 2\sigma^2}} \ , \quad (41)$$

$$CNR = \frac{2\nu}{c^2 \sigma^2} \ . \quad (42)$$

By estimating ν using (34) an effective value ν_{eff} will be obtained, related to the true value by [10]

$$\nu_{eff} = \nu\left(1+\frac{1}{CNR}\right)^2. \tag{43}$$

When using the value ν_{eff} instead of ν the distribution is a reasonable fit to the tail of the data with added noise, but it gives a poor fit for the low amplitude values. Note, though, that it is the tail region that is responsible for such an important detection characteristic as probability of false alarm. Hence, according to Watts, for single pulse detection in presence of thermal noise as well as sea clutter, a reasonable guide to detection performance may be achieved by treating the interference signal as being K-distributed but with a modified shape parameter ν_{eff}. For single-pulse ideal CFAR detection with CNRs of 0 dB or less, a good estimate of performance will generally be obtained by assuming Rayleigh statistics.

2.3.2.2.3 First and second sample moments based method

Alternatively, estimates of the shape and scale parameters of the K-distribution may be obtained by using the first and second moments [4,14]. The ratio of the second sample moment to the squared sample mean, given by

$$\frac{m_2}{m_1^2} = \frac{4\nu[\Gamma(\nu)]^2}{\pi[\Gamma(\nu+0.5)]^2}, \tag{44}$$

may be solved numerically to estimate the shape parameter ν. Although this method does not have a closed-form solution, the smaller variability in the lower-order moments usually yields better results than higher moment methods. The scale parameter c may be estimated from the first moment, that is

$$c = \frac{\Gamma(\nu+0.5)\sqrt{\pi}}{\Gamma(\nu)m_1}. \tag{45}$$

As mentioned before, the moments based estimation of the K-distribution parameters does not perform well when the available number of samples is limited because in this case the sample moments do not represent good estimates of the true moments, and meaningful parameter estimates of the distribution are not obtained.

The moments based methods of estimating of the K-distribution parameters are also particularly susceptible to test data sets in which a target is unwittingly present. In this case they provide an estimate of the shape parameter which is much less than the real value of this parameter.

2.3.2.3 Arithmetic and geometric means based (Raghavan's) method

A method proposed by Raghavan [4,14] was found to be less susceptible to these problems within any particular test sample. In this method the similarity between the K-distribution and the simpler Gamma distribution given below is exploited. The PDF for the Gamma distribution is given by the expression:

$$f_G(a) = \frac{a^{\beta-1}}{b^{\beta}\Gamma(\beta)}\exp(-\frac{a}{b}) \quad . \tag{46}$$

where β and b are the shape and scale parameters of the Gamma distribution, respectively. It is straight forward to derive relationships between the arithmetic and geometric means of the sample distribution and the ML estimates of the shape and scale parameters of the corresponding Gamma distribution. Specifically, if the arithmetic and geometric means of the sample distribution, m_a and m_g, are derived by expressions:

$$m_a = \frac{1}{n}(a_1+...+a_n) = \frac{1}{n}\sum_{i=1}^{n} a_i \quad , \tag{47}$$

$$m_g = \sqrt[n]{(a_1...a_n)} = (\prod_{i=1}^{n} a_i)^{1/n} \quad , \tag{48}$$

then the ML estimates of the shape and scale parameters of the corresponding Gamma distribution, β and b, for the sample are related by the following equations [4]:

$$p_n = \frac{m_a}{m_g} = \beta\exp(-\psi(\beta)) \quad , \tag{49}$$

$$b = \frac{m_a}{\beta} \quad , \tag{50}$$

where $\psi(\beta)$ is the Digamma function which is defined as $\psi(x) = \frac{d}{dx}[\ln(\Gamma(x)] = \frac{\Gamma'(x)}{\Gamma(x)}$.

The function on the right-hand side of (49) may be computed easily by the following expression [4]:

$$x\exp[\psi(x)] = \prod_{k=0}^{\infty}\left(1+\frac{1}{x+k}\right)^{-1}\exp\left(\frac{1}{x+k}\right) \quad . \tag{51}$$

By equating the first and second moments of the random values, one of which is K-distributed and another has a corresponding Gamma distribution, the parameters ν and c for the K-distribution may be expressed as functions of the Gamma density parameters (β and b). Specifically, the parameter β of the equivalent Gamma density may be expressed in terms of the parameter, ν, by the following relation:

$$\beta = \left[\frac{4\nu\Gamma^2(\nu)}{\pi\Gamma^2(\nu+0.5)-1} \right]^{-1}. \tag{52}$$

Raghavan argues that because the K- and Gamma distributions are similar over a range of parameter values where the K-distribution is highly non-Rayleigh, an estimate that approximates the ML estimate of the K-distribution can be delivered with the same statistic used to compute the ML estimate for the Gamma distribution. Therefore, a good estimate for the shape parameter ν of the K-distribution can be based upon the relation $p_n = m_a/m_g$. Since m_a, m_g and p_n are evaluated from a finite number of samples, these quantities are in general random values and their statistics depend upon the PDF of the samples $\{a_i\}$ and on the number of samples n, used in the computation. Raghavan derives the expected value of p_n for the K-distribution in terms of ν :

$$E(p_n) = \frac{\Gamma(0.5+\nu-0.5n^{-1})\Gamma(1.5-0.5n^{-1})}{\Gamma(\nu)} \left[\frac{\Gamma(\nu-0.5n^{-1})\Gamma(1-0.5n^{-1})}{\Gamma(\nu)} \right]^{n-1} \tag{53}$$

for $n > (2\nu)^{-1}$.

Using the observed value of p_n, this equation can be solved numerically in order to estimate the shape parameter ν of the K-distribution. Then using the expression for the first order moment of the K-distribution random values, the scale parameter c may be obtained from the estimate $\hat{\nu}$ and m_a as

$$\hat{c} = \frac{2}{m_a} \frac{\Gamma(\hat{\nu}+0.5)\Gamma(1.5)}{\Gamma(\hat{\nu})}. \tag{54}$$

The K-distribution with the derived parameters (ν and c) has approximately the same arithmetic and geometric means as the original test set. This method is a numerically fast method of estimating the ML parameters at the cost of greater complexity when compared to the methods of moments approach.

For highly and moderately non-Rayleigh data, Raghavan's method gives better estimates for small sample sizes in comparison to the method of moments techniques. Unfortunately, this method also has limitations. When the K-distribution is more nearly Rayleigh, there is little similarity between the K- and Gamma distributions. The

applicability of Raghavan's method is diminished in such cases. A comparison of higher order moments of a K-distributed random variate and those of the equivalent Gamma variate shows that this approach yields good approximations to the K-distribution only for the range $0.2 < \nu \leq 2$, which physically represents spiky clutter characteristics.

2.3.2.4 Modified chi-square test

It was shown in [3], that from the radar perspective, standard statistical tests such as the chi-square goodness of fit test are of limited use for clutter data. The reason is that these tests place an equal importance in the goodness of fit on all regions of the probability space. In radar applications, goodness of fit of clutter statistics to a model is important mainly in the low PFA region.

To obtain a procedure which brings out the relative results of various models in the important low PFA region, the modified chi-square test was proposed [3]. Boundaries of intervals in this test were determined for each of the statistical models in the amplitude region for which PFA is ≤ 0.1, assigning a zero weighting in the amplitude region where the PFA is greater than 0.1.

The modified χ_m^2 index is defined as

$$\chi_m^2 = \sum_{i=1}^{K} \frac{[f_i - N(0.1p_i)]^2}{N(0.1p_i)}, \tag{55}$$

where K is the number of intervals into which the low PFA region of the statistical model is divided, f_i is the observed number of occurrences of a clutter sample having an amplitude within the i^{th} interval, N is the total number of amplitude samples forming the histogram, and $N(0.1p_i)$ is the weighted expected number of occurrences in the i^{th} interval for the statistical model in the low PFA region.

The ensemble average of the modified chi-squared index χ_m^2 for a group of resolution cells can be computed for each model and compared to give a quantitative measure of the relative goodness of fit in the low PFA region. A lower modified chi-squared index χ_m^2 value indicates a better fit of the model to the data being considered.

2.3.2.5 The choice of the method for estimating parameters of the K-distribution

The discussion of the previous sections of the methods of estimation of the K-distribution parameters, may be summarised as follows:

- the ML method theoretically provides the best statistical estimates of the parameters of the K-distribution but this method is impractical from a computational point of view;

- the moments based methods guarantee good quality of the estimates of the parameters of the K-distribution for large sample sizes ($n > 1000$) when there is no unwitting presence of a target;

- the moments based methods are better suited to the estimation of large values of the shape parameter of the K-distribution. This can probably be attributed to the increased variability in the moments of the heavier tailed distributions that occur for small values of the shape parameter;

- the first and second sample moments based method (FSM) is more accurate than Watts's method (using the second and fourth sample moments) because of the smaller variability in the lower-order sample moments;

- Raghavan's method works well even for small sample sizes if the value of the shape parameter of the K-distribution is small ($\nu < 2$) where the correspondence between the K- and Gamma distributions is the strongest.

- if the observed data has a low clutter-to-noise ratio (CNR < 10 dB), then the resulting amplitude distribution will be significantly different from a standard K-distribution and the interference signal should be treated as being K-distributed but with the modified shape parameter ν_{eff}.

Comparison of Raghavan's method and the moments based methods shows that each of them can perform well and provide accurate estimates of the parameters of the K-distribution for specific conditions. With sufficient computational resources and time, it is possible to calculate three pairs of the estimates for the shape and scale parameters of the K-distribution (using FSM, Watts's and Raghavan's methods) and then choose the best pair for each particular situation by the following simple logic:

- if the observed data has a low clutter-to-noise ratio (CNR < 10 dB) then the pair of estimates obtained by using the modified Watts's method should be chosen;

- if the observed data has a high clutter-to-noise ratio (CNR > 10 dB) and the sample size is big enough, or Raghavan's method gives the estimate for the shape parameter $\hat{\nu} > 2$, then the pair of estimates obtained by using the FSM method should be chosen;

- otherwise the pair of estimates of the shape and scale parameters of the K-distribution obtained by using the Raghavan's method should be chosen.

This logic is determined by the fact that CFAR processors are much more sensitive to errors in the estimated value of the shape parameter for small values of this parameter than for large values [22]. The accuracy of the estimates of the shape parameter of the K-distribution is more important for smaller values of this parameter. For large values of the shape parameter, the K-distribution approaches the Rayleigh distribution and its shape becomes insensitive to this parameter. Thus the number of samples needed to estimate the shape parameter to a given accuracy increases rapidly with increasing values of this parameter, but the number of samples needed to obtain a PDF that fits the data to a given level of the mean-square difference between the true PDF and the PDF given by the estimated parameters decreases with increasing the value of the shape parameter [14].

Another simple and effective way to choose the best estimates of the K-distribution parameters for the particular sea clutter data set is to calculate the modified chi-squared index χ_m^2 values for the statistical models, parameters of which are defined by using the considered estimation methods (Watts's, FSM and Raghavan's), and compare these values. The model with the best pair of estimates gives the lowest value of the modified chi-squared index χ_m^2.

3. Performance results

In order to illustrate the advantages of using the compound K-distribution model over other models for estimating the parameters of sea clutter, and compare the different methods for estimating of the K-distribution shape and scale parameters, different models for the amplitude distribution have been applied to a single data set from database of experimentally collected sea clutter using the INGARA radar.

The INGARA radar system was developed within the Tactical Surveillance Systems Division of DSTO as a technology demonstrator aimed at investigating and demonstrating the application of Synthetic Aperture Radar to the unique surveillance challenges posed by the large sparsely populated areas across the Northern Australian coastline. The main sensor of the INGARA system is a coherent, horizontal polarised, X-band multi-mode radar system. The flexible nature of the design of this radar system, based on open architectures, has allowed for the addition of maritime surveillance modes specifically designed to collect radar backscatter from the surface of the ocean (sea clutter).

The data set, that has been chosen, was taken from the database of experimental data, which were collected between 9 and 11 November 1993 at Port Noarlunga South, South Australia, using the INGARA radar system at frequency of 9.375 GHz . For this trial the radar was set up on a cliff-top approximately 30 m above the sea surface. Most of the data consists of files containing 30 second of data collected with the radar

pointed in 15^0 intervals in azimuth angle between each data run. The azimuth angular range of the measurements was from 210^0 to 345^0.

The chosen data set consists the data for the antenna looking at an angle of 315^0. The measured meteorological conditions were :

- Wind Speed 5.6 m/s
- Wind Direction 138^0
- Air Temperature 15.8^0C
- Sea Temperature 19.4^0C
- Relative Humidity 54.7
- Barometric Pressure 1015

Observations of the sea made from the radar site note that it was slight, with small wind ripples, probably sea state 2.

The INGARA radar coherently sampled data at a 50 MHz rate (3 m in range). Each sample is a 4-byte word that represents the HI and HQ returned signals at a particular sample instant (ie. each channel is digitised to 8 bits). The symbols HI and HQ refer to the horizontal polarised inphase and quadrature channels, respectively.

The main parameter of interest to radar designer is the false alarm rate, which sets the detection threshold for the radar system. It is (usually) determined through fitting the assumed model for the clutter to a limited sample of clutter returns, and deducing the threshold from the statistics associated with the model. In this study, because we have a sufficiently large data set, we can make a direct comparison of the false alarm rates experienced with the different thresholds associated with the models, to assess the relative merits of the models. In general, the better the fit of the clutter parameters to the actual data, the closer will be the desired and actual false alarm rates.

An image of the amplitude of the clutter for the data set used in the analysis is presented on Figure 1. Each range window of the image represents 300 consecutive range samples, or 900m span, and the time duration is approximately 0.3 s with a PRF of 333.3 Hz. The distance from the radar to the first range bin is 3384 m. The data has been normalised to have unit second sample moment.

Figure 2 presents the averaged temporal (for the speckle component) and spatial (for the Gamma component) autocorrelation functions for the data set. At any range, the return fluctuates with a time constant of approximately 10 ms as the scatterers within the patch move with the internal motion of the sea and change their phase relations. The local mean level varies with the range owing to bunching of the scatterers. The correlation length of the sea surface in the range direction is about 6.5 m. The duration (0.3 s) of the data collection is not sufficient for the bunching to change at any given range.

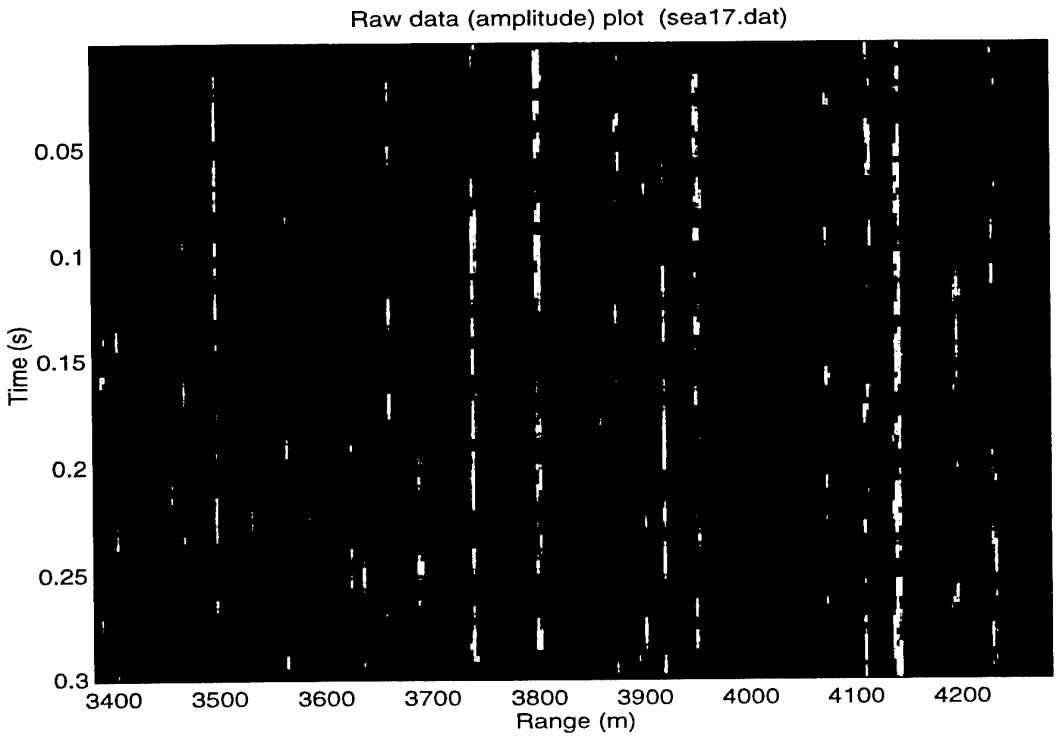

Figure 1 Experimentally collected sea clutter

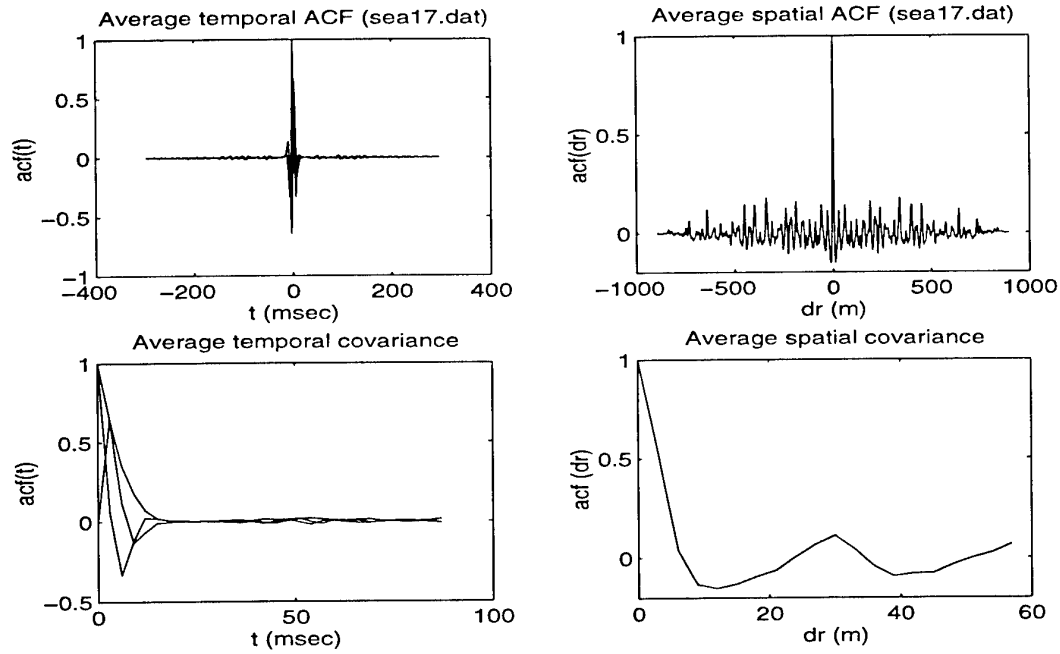

Figure 2 Temporal and spatial autocorrelation functions for experimentally collected sea clutter

As the sea clutter shows two dominant fluctuation components, the data set contains many more independent values of the speckle than of the underlying modulation. The analysis has therefore been adapted to assume these two components, with a large number of independent speckle samples and a limited number of the modulation.

Subsampling of the data was done to remove any correlations resulting from underlying cross-section modulations before parameter estimation was performed. In particular, statistically independent samples of the sea clutter amplitude distribution were obtained by taking echo samples from resolution cells that are separated by distances greater than the decorrelation length of the large scale effects for all realisations of received signals (each realisation presents the echo signals from 300 range bins).

For each realisation the following parameters were calculated:

- estimates of the parameters for the considered distributions (Log-Normal, Weibull and K-distribution) using different methods;
- the ratio of the theoretical to the observed moments for the first six moments for each combination of parameter estimation and distribution considered;
- the modified chi-squared index χ^2_m values for each combination of parameter estimation and distribution considered;
- standard deviation σ_ν of the estimates of the K-distribution shape parameter by each method.

As each calculated parameter was averaged through all realisations, the presented results are the mean values of these parameters.

Table 1 lists estimates of the parameters for the K-distribution using the three methods (Watts's, first and second moments (FSM) and Raghavan's) as well as estimates of the parameters for the Log-Normal and Weibull distributions using the ML method, and for the Weibull distribution using Menon's method.

Table 1 Estimates of the distribution parameters by different methods for experimentally collected sea clutter data set

Parameter	LN	W(ML)	W(Menon)	K(R)	K(W)	K(FSM)
ν	(μ) -0.5645	(γ) 1.3351	(γ) 1.6254	1.3131	0.6160	0.8256
c	(σ^2) 0.6297	(ϖ) 0.8366	(ϖ) 0. 8128	2.4039	1.5498	1.8075

Note, that all the estimators of the K-distribution parameters are constrained so that the maximum value of the shape parameter ν they estimate is 50. At ν=50 the distribution is effectively Rayleigh, and in practice the estimated value of the shape

parameter rarely exceeds 20. Thus, limiting ν does not significantly change the mean-square difference between the true PDF and the PDF given by the estimated parameters, but such a limit does allow an easier comparison of standard deviations for different estimation techniques [14]. This upper limit is used in the numerical implementations of the FSM and Raghavan's estimators to obtain a finite search area and to ensure convergence. Watts's method sometimes yields negative values of ν; when this happens, the estimator outputs ν equal to 50.

The estimated parameters were used to produce the probability density functions for all the considered distributions, and to compare the results with the amplitude histogram of the sea clutter data.

Figure 3 is a plot of the probability density functions for the K-distribution, estimated by Watts's, first and second moments (FSM) and Raghavan's methods.

Figure 4 is a plot of the probability density functions for the Log-Normal and Weibull distributions estimated by the maximum likelihood (ML) method and the K-distribution estimated by Watts's method.

For comparison purposes K-distribution parameters using the ML method were calculated. Figure 5 is a plot of the probability density function for the K-distribution estimated using this procedure.

According to the presented results the analysis of amplitude statistic of the clutter has demonstrated that the data distribution develops a longer tail than the Rayleigh distribution. The estimated values of the shape parameters for the Weibull and the K-distributions correspond to those for the spiky clutter.

Menon's method for the Weibull distribution overestimates the value of the shape parameter compared to the optimal ML method, and as a result the threshold might be set too low. The consequence of this would be the performance degradation due to a notable increase in probability of false alarm (PFA).

The different methods for estimation of the K-distribution parameters produce quite different estimates of the parameters. In particular, the moments methods estimate the value of the shape parameter to be much lower than under ML or Raghavan's method. The reasons for this might be not only the different approaches used by these methods, but the conditions in which the data set was collected.

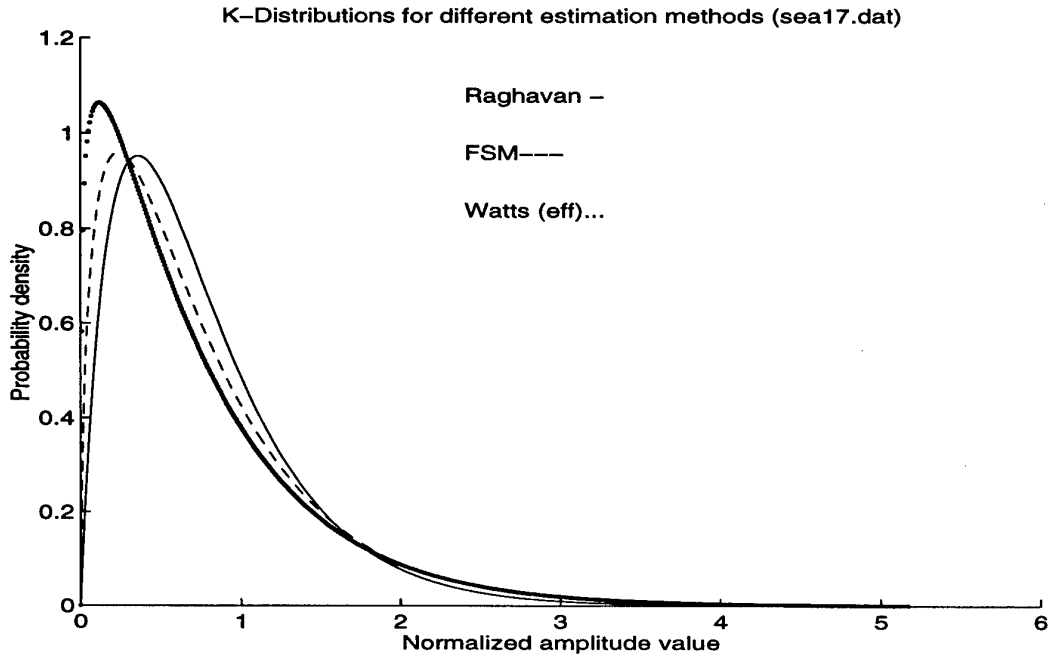

Figure 3 K-PDFs estimated by different methods for experimentally collected sea clutter

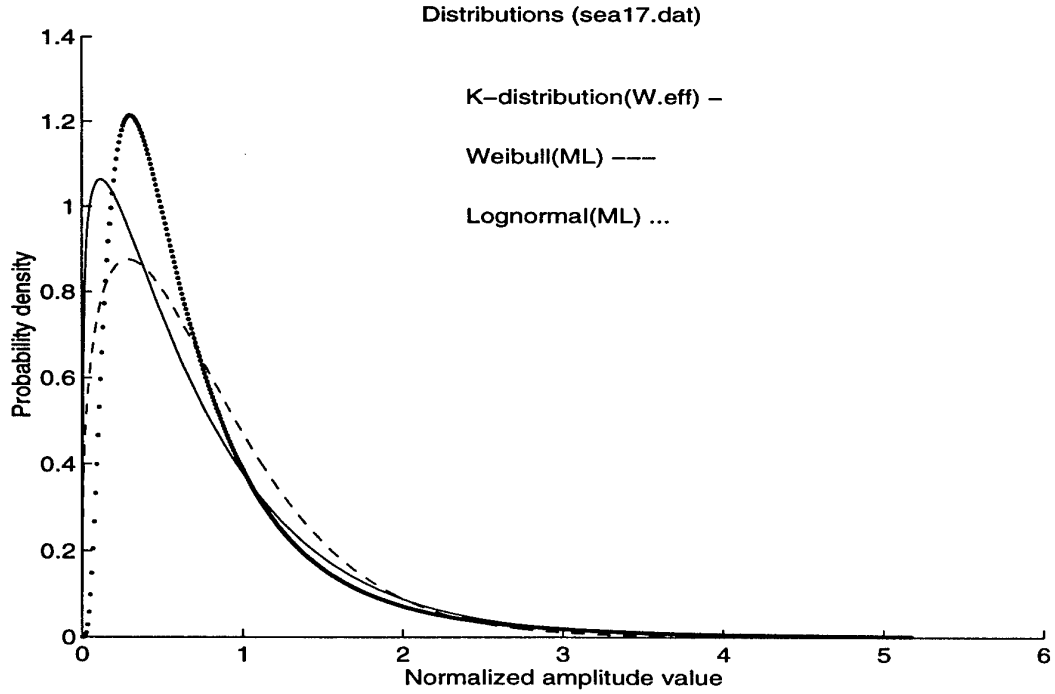

Figure 4 Different estimated PDFs for experimentally collected sea clutter

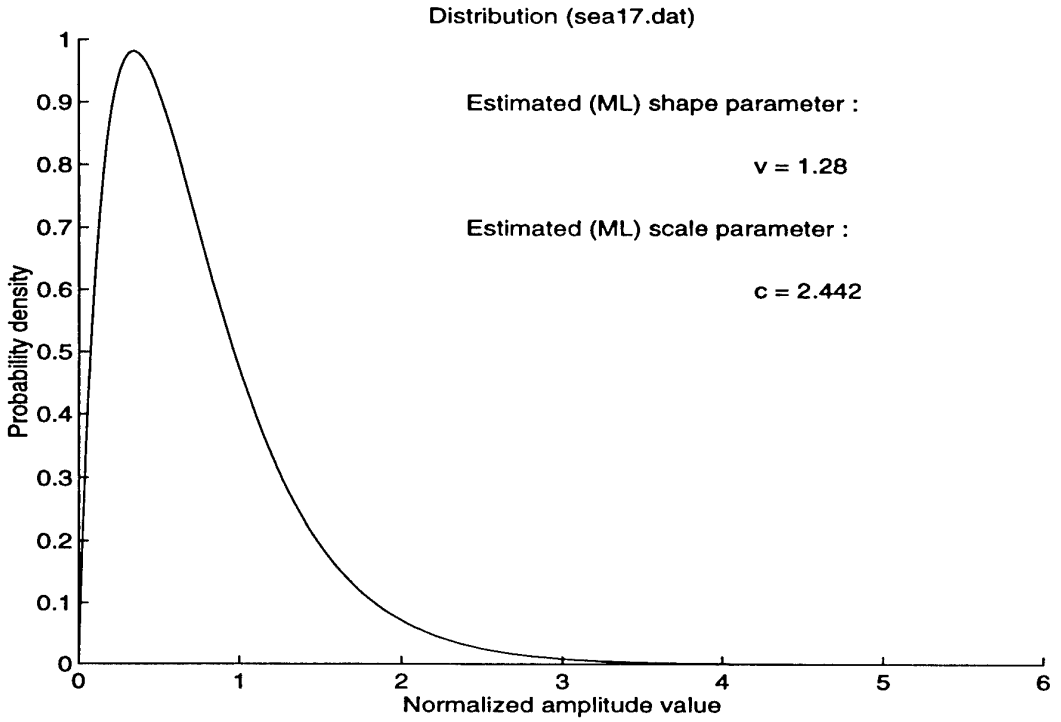

Figure 5 K-PDF estimated by ML method for experimentally collected sea clutter

The data set used in the analysis has a low clutter-to-noise ratio (CNR) . This should have a strong influence on the performance results of the K-distribution analysis, because the presence of thermal noise, which cannot be neglected, modifies the original distribution of the sea clutter. The low amplitude values of the sea clutter distribution are the most affected by the noise: for small values of the amplitude, the additive noise increases the final power significantly, whereas for large values it leaves the power almost unchanged. As a result, the distribution of the sea clutter combined with additive noise is not K-distributed, and a closed-form expression defining this distribution does not exist. For large values of the shape parameter ν, when the amplitude is concentrated about its mean value with a low probability for small values of the amplitude, the added noise makes little difference to the final distribution. However, for small values of the shape parameter ν the amplitude distribution is very spiky with a high probability for small values of the amplitude. These are affected considerably by additive noise.

As the ML method of estimation uses the suggestion that experimentally collected sea clutter data is K-distributed and the expected value of the shape parameter is small ($\nu<2$), the results are quite predictable: the K-distribution with the parameters under the ML estimation technique in this case should fit well the lower amplitude area of the experimentally collected data histogram, which is mainly distorted by the noise, but give poor results in the tail area.

Raghavan's method uses the similarity between the K-distribution and the simpler Gamma distribution in order to deliver an estimate of the K-distribution shape parameter, that approximates the ML estimate, with the same statistic used to compute the ML estimate for the Gamma distribution shape parameter. Because this method based on the assumption that experimentally collected sea clutter data is K-distributed, it suffers from the same problem as the original ML method for the K-distribution in low CNR environment: the K-distribution with the parameters under Raghavan's estimation technique fits well the lower amplitude area of the experimentally collected data histogram, which is mainly distorted by the noise, but gives poor results in the tail area.

In the contrast, the moments based methods provide with a reasonable fit to the tail of the data with added noise, but give a poor fit for the low amplitude values. The moments methods use for estimation the observed moments, which are less distorted by noise than the amplitude histogram itself, especially the higher order moments. Because the tail region is responsible for such an important detection characteristic as probability of false alarm, from the point of view of practical radar applications the moments methods are better in low CNR situations.

According to published results [10], the ideal CFAR performance is more sensitive to the presence of thermal noise, approaching more rapidly the detection performance expected in noise alone as the clutter-to-noise ratio falls. For fixed threshold detection (considered in this section) the performance in spiky clutter is determined to a large extent by the clutter spikes even for a relatively low clutter-to-noise ratio.

To provide a comparison of the different models, the ratio of the theoretical to the observed moments for the first to sixth moments for each combination of parameter estimation and distribution considered were calculated. Table 2 lists the results of these calculations.

The result for the Log-Normal distribution suggests overestimating of the tail length of the experimental data amplitude histogram. The consequence of using the estimated parameters of the Log-Normal distribution for the threshold setting would be increased detection losses, as the threshold would be set too high. Thus, poor results should be expected in this case for small and middle size target detection situations.

Table 2 Ratio of the theoretical and observed moments for different estimation methods for experimentally collected sea clutter data set

Th./Ob.	LN	W(ML)	W(Menon)	K(R)	K(W)	K(FSM)
mt1/mo1	1.0212	1.0078	0.9540	1.0000	0.9489	1.0000
mt2/mo2	1.1466	0.9318	0.7444	0.9000	1.0000	0.9998
mt3/mo3	1.6144	0.7401	0.4721	0.7123	1.0025	0.9172
mt4/mo4	3.54 00	0.5508	0.2719	0.5460	1.0000	0.8294
mt5/mo5	15.01 00	0.4206	0.1574	0.4439	1.0639	0.8020
mt6/mo6	147.5200	0.3445	0.0964	0.3995	1.2516	0.8625

The Weibull distribution under the ML method and K-distribution under Raghavan's method are similar. The lower moments are in a good agreement with real data , but the higher moments are underestimated. Thus, the distributions with these parameters give a good fit to lower amplitude area , but they have shorter tails compare to experimental data histogram. The result of usage of the parameters of such distributions for threshold setting would be an increased PFA.

Menon's approach for the Weibull distribution fares poorly. This method gives the worst results of the techniques we have investigated.

The second and fourth moments under Watts's method are forced to unity, and this method has the best higher moments ratio amongst all the combinations considered. The FSM method lies roughly between the Raghavan's method and the Watts's method.

To verify the results of the amplitude fitting a modified version of the Chi-squared statistical test was used. Table 3 lists the modified chi-squared index χ_m^2 values for all these distributions, and the standard deviation σ_v of the estimates of the K-distribution shape parameter by each method.

Table 3 Modified chi-squared index χ_m^2 values and standard deviation σ_v for different estimation methods for experimentally collected sea clutter data set

Parameter	LN	W(ML)	W(Menon)	K(R)	K(W)	K(FSM)
χ_m^2	709.7961	202.5867	73762.2000	204.8008	111.0911	121.6138
σ_v				0.3109	0.2136	0.1986

The comparison of the modified chi-squared index χ_m^2 values for all the distributions shows, that the best results in the important tail area can be achieved by applying the K-distribution model to the sea clutter. Among the K-distribution parameters estimation methods, the distribution with the parameters using Watts's method gives a better fit in this region to the experimentally collected data histogram than the others.

Analysis of the standard deviation σ_v of the estimates of the K-distribution shape parameter by each method shows that the FSM method has a smaller deviation than Watts's method, which can be explained by smaller variability in the lower-order sample moments. Quite a big value of the standard deviation for Raghavan's method follows from the fact that it gives a good fit to the lower amplitude values, which are the most distorted by noise.

To obtain the number of false alarms for the analysing data set, the thresholds were set, derived for the Log-Normal, Weibull distributions and K-distribution under ML

estimation, and for the K-distribution under the all three other considered estimation techniques for probability of false alarm of 10^{-3} and 10^{-4}.

The data set corresponds to 100 pulses by 300 range bins, totalling 30,000 sample points. Hence, on average 30 false alarms would be expected for a false alarm rate of 10^{-3} (ie one false alarm on every thousand samples), if the underlying distribution is correct and the samples are independent. Similarly a false alarm rate of 10^{-4} corresponds to 3 false alarm on average. The spatial correlation of the local variations in clutter mean level may result in a similar 'bunching' of false alarms.

Table 4 displays the number of amplitude values of the sample lying above the thresholds corresponding to false alarm rates 10^{-3} and 10^{-4} respectively.

Table 4 Number of false alarms for different estimation methods for experimentally collected sea clutter data set

PFA	LN	W(ML)	K(ML)	K(R)	K(W)	K(FSM)
10^{-3}	0	305	123	129	25	52
10^{-4}	0	40	14	16	0	0

According to the expectations, the Log-Normal distribution is over-conservative. As has been reported by other authors, in the tail region the K-distribution lies between the Log-Normal distribution and the Weibull distribution. The Weibull distribution yields very poor results. The K-distributions under ML estimation and Raghavan's method are better. Watts's method is by far the best method, and the FSM method lies between Raghavan's methods and Watts's method.

As was mentioned before, the presence of quite strong thermal noise (the experimentally collected data set has a low clutter-to-noise ratio) distorts the resulting amplitude distribution from the assumed noise-free K-distributed clutter, especially at the low amplitude values. Raghavan's and the FSM methods are sensitive to the lower moments of the amplitude distribution. However Watts's method relies on the second and fourth moments for estimation of the effective value of the shape parameter. The higher moments are less affected by distortions to the lower amplitude values of the clutter amplitude distribution by the noise. This indicates why for this data set, Raghavan's and the FSM methods tend to fit the lower amplitude values region better, but Watts's method is better in modelling the false alarm rates.

Note, that some influence of the antenna beam shape changes the statistics of experimentally collected sea clutter as well. This might be one of the reasons why the modified Watts's method may be slightly conservative. Another reason is sensitivity of the estimation accuracy for higher order moments to the limited sample size.

4. Summary

In this report the three main distributions for modelling of sea clutter (Log-Normal, Weibull and K-distribution) have been briefly reviewed. We have compared the methods against a single data set obtained from the data base of recording signals of the cliff-top positioned INGARA system. It has been shown that the K-distribution is the most promising of the models, especially for estimating thresholds for a low probability of false alarm.

Several existing methods for estimation of the parameters of the K-distribution have been analysed and some recommendations about their implementation have been suggested.

Further work needs to be done in this area, to investigate the effect of radar system characteristics, look direction, polarisation and area conditions on the parameters of the clutter distribution. In order that the models used for sea clutter are validated, it is necessary to have a good amount of experimentally collected clutter data sets which respond to following requirements:

- each data set must include several hundreds returned pulses from every range bin for the purpose of removing of the speckle component by averaging when estimating the correlation length of the large scale effects;

- each data set must include echo signals from at least several hundred (preferably thousand) range bins in order to have enough independent samples for analysis after subsampling to remove data correlations;

- sea clutter data records must be of good quality (i.e. have high clutter-to-noise ratio) to minimise the effect of noise on clutter characteristics;

- the antenna beam shape has to be designed to minimise its influence on the recorded data.

The validation of these models will improve the accuracy of the detection performance prediction for maritime surveillance radar systems.

5. References

1. Hou, X.-Y., and Morinaga, N. "Detection performance in K-distributed and correlated Rayleigh clutters", IEEE Trans. AES, Vol.25, No 5, Sept. 1989, pp. 634-641.

2. Nathanson, F.E., Reilly, J.P., and Cohen, M.N. "Radar design principles: signal processing and the environment", 2nd ed., McGraw-Hill inc., 1990.

3. Chan, H.C. "Radar sea clutter at low grazing angles", IEE Proc., Vol. 137, Pt. F, No 2, Apr. 1990, pp. 102-112.

4. Raghavan, R.S. "A method for estimating parameters of K-distributed clutter", IEEE Trans. AES, Vol. 27, No 2, March 1991, pp. 238-246.

5. Ward, K.D., Baker, C.J., and Watts, S. "Maritime surveillance radar. Part 1: Radar scattering from the ocean surface", IEE Proc., Vol. 137, Pt. F, No 2, Apr. 1990, pp. 51-62.

6. Ward, K.D., Baker, C.J., and Watts, S. "Maritime surveillance radar. Part 2: Detection performance prediction in sea clutter", IEE Proc., Vol. 137, Pt. F, No 2, Apr. 1990, pp. 63-72.

7. Watts, S. " A practical approach to the prediction and assessment of radar performance in sea clutter", IEEE International Radar Conference, 1995, pp. 181-186.

8. Ward, K.D. "A radar sea clutter model and its application to performance assessment", Radar-82, pp. 203-207.

9. Watts, S. "Radar detection prediction in sea clutter using the compound K-distribution model", IEE Proc., Vol. 132, Pt. F, No 2, 1985, pp. 613-620.

10. Watts, S. "Radar detection prediction in K-distributed sea clutter and thermal noise", IEEE Trans. AES, Vol. 23, No 1, Jan 1987, pp. 40-45.

11. Farina, A., Russo, A., and Studer, F.A. " Coherent radar detection in log-normal clutter", IEE Proc., Vol. 133, Pt. F, No 1, Feb. 1986, pp. 39-54.

12. Conte, E., and Longo, M. "On a coherent model for log-normal clutter", IEE Proc., Vol. 134, Pt. F, No 2, Apr. 1987, pp. 198-201.

13. Haykin, S. "Adaptive filter theory", 2nd ed., Prentice Hall, 1991.

14. Joughin, I.R., Percival, D.B., and Winebrenner, D.P. "Maximum likelihood estimation of K-distribution parameters for SAR data", IEEE Trans. Geosci. Remote Sensing, Vol. 31, 1993, pp. 989-999.

15. Schleher, D.C. "Radar detection in Weibull clutter", IEEE Trans. AES, Vol. 12, No 6, Nov. 1976, pp. 736-743.

16. Farina, A., Russo, A., Scannapieco, F., and Barbarossa, F. "Theory of radar detection in coherent Weibull clutter", IEE Proc., Vol. 134, Pt. F, No 2, Apr. 1987, pp. 174-190.

17. Rifkin, R. "Analysis of CFAR performance in Weibull clutter", IEEE Trans. AES, Vol. 30, No 2, Apr. 1994, pp. 315-328.

18. Ravid, R., Levanon, N. "Maximum-likelihood CFAR for Weibull background", IEE Proc., Vol. 139, Pt. F, No 3, June 1992, pp. 256-264.

19. Menon, M.V. "Estimation of the shape and scale parameters of the Weibull distribution", Technometrics, No 15, 1963,pp. 175-182.

20. Baker, C.J. "K-distributed coherent sea clutter", IEE Proc., Vol. 138, Pt. F, No 2, Apr. 1991, pp. 89-92.

21. Nohara, T.J., and Haykin, S. "Canadian East Coast radar trial and the K-distribution", IEE Proc., Vol. 138, Pt. F, No 2, Apr. 1991, pp. 80-88.

22. Armstrong, B.C., and Griffiths, H.D. "CFAR detection of fluctuating targets in spatially correlated K-distributed clutter", IEE Proc., Vol. 138, Pt. F, No 2, Apr. 1991, pp. 139-152.

DISTRIBUTION LIST

Analysis of Sea Clutter Data

Irina Antipov

AUSTRALIA

DEFENCE ORGANISATION

Task Sponsor: DGAD

S&T Program

Chief Defence Scientist
FAS Science Policy
AS Science Corporate Management
} shared copy

Director General Science Policy Development
Counsellor Defence Science, London (Doc Data Sheet)
Counsellor Defence Science, Washington (Doc Data Sheet)
Scientific Adviser to MRDC Thailand (Doc Data Sheet)
Director General Scientific Advisers and Trials/Scientific Adviser Policy and Command (shared copy)
Navy Scientific Adviser (Doc Data Sheet and distribution list only)
Scientific Adviser - Army (Doc Data Sheet and distribution list only)
Air Force Scientific Adviser
Director Trials

Aeronautical and Maritime Research Laboratory
Director
Chief of Air Operations Division Dr. C. Guy

Electronics and Surveillance Research Laboratory
Director
Chief of Wide Area Surveillance Division Dr. D.H. Sinnott
Chief of Electronic Warfare Division Dr. M.L. Lees
Chief of Tactical Surveillance System Division Dr. D. Gambling
Research Leader Tactical Surveillance Dr. D. Heilbronn
Head of Sensor Processing, TSSD, Dr. J. Whitrow
Task Manager B. Reid, TSSD
Author: I. Antipov, TSSD

DSTO Library
Library Fishermens Bend
Library Maribyrnong
Library Salisbury (2 copies)
Australian Archives
Library, MOD, Pyrmont (Doc Data sheet only)

Capability Development Division

Director General Maritime Development (Doc Data Sheet only)
Director General Land Development (Doc Data Sheet only)
Director General C3I Development (Doc Data Sheet only)

Navy

SO (Science), Director of Naval Warfare, Maritime Headquarters Annex, Garden Island, NSW 2000. (Doc Data Sheet only)

Army

ABCA Office, G-1-34, Russell Offices, Canberra (4 copies)
SO (Science), DJFHQ(L), MILPO Enoggera, Queensland 4051 (Doc Data Sheet only)
NAPOC QWG Engineer NBCD c/- DENGRS-A, HQ Engineer Centre Liverpool Military Area, NSW 2174 (Doc Data Sheet only)

Intelligence Program

DGSTA Defence Intelligence Organisation

Corporate Support Program (libraries)

OIC TRS, Defence Regional Library, Canberra
Officer in Charge, Document Exchange Centre (DEC), 1 copy
*US Defence Technical Information Center, 2 copies
*UK Defence Research Information Centre, 2 copies
*Canada Defence Scientific Information Service, 1 copy
*NZ Defence Information Centre, 1 copy
National Library of Australia, 1 copy

UNIVERSITIES AND COLLEGES

Australian Defence Force Academy
Library
Head of Aerospace and Mechanical Engineering
Deakin University, Serials Section (M list), Deakin University Library, Geelong, 3217 Senior Librarian, Hargrave Library, Monash University
Librarian, Flinders University

OTHER ORGANISATIONS

NASA (Canberra)
AGPS
State Library of South Australia
Parliamentary Library, South Australia

OUTSIDE AUSTRALIA

ABSTRACTING AND INFORMATION ORGANISATIONS

INSPEC: Acquisitions Section Institution of Electrical Engineers
Engineering Societies Library, US
Documents Librarian, The Center for Research Libraries, US

INFORMATION EXCHANGE AGREEMENT PARTNERS

Acquisitions Unit, Science Reference and Information Service, UK
Library - Exchange Desk, National Institute of Standards and Technology, US
National Aerospace Laboratory, Japan
National Aerospace Laboratory, Netherlands

SPARES (10 copies)

Total number of copies: 40

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
DOCUMENT CONTROL DATA

1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE

Analysis of Sea Clutter Data

3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)

Document	(U)
Title	(U)
Abstract	(U)

4. AUTHOR(S)

Irina Antipov

5. CORPORATE AUTHOR

Electronic and Surveillance Research Laboratory

PO Box 1500
Salisbury South Australia 5108

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-TR-0647	AR-010-489	Technical Report	March 1998

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
Z9505/13/161	ADA 95/080	DGAD	32	22

13. DOWNGRADING/DELIMITING INSTRUCTIONS

Not Applicable

14. RELEASE AUTHORITY
Chief, Tactical Surveillance System Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE CENTRE, DIS NETWORK OFFICE, DEPT OF DEFENCE, CAMPBELL PARK OFFICES, CANBERRA ACT 2600

16. DELIBERATE ANNOUNCEMENT

No limitations

17. CASUAL ANNOUNCEMENT No limitations

18. DEFTEST DESCRIPTORS

Sea clutter, surveillance radar, ocean surveillance, P-3C aircraft

19. ABSTRACT

This report presents the results of comparative analysis of the models which have been applied to sea clutter amplitude distribution, and the methods for estimation of their parameters. More detail consideration among others is given to the K-distribution as the most appropriate model for sea clutter in the low Probability of False Alarm (PFA) region.

VS

[redacted]

TECHNICAL REPORT DSTO-TR-0647 AR-010-489 MARCH 1998

ELECTRONICS AND SURVEILLANCE RESEARCH LABORATORY
PO BOX 1500 SALISBURY SOUTH AUSTRALIA, 5108
AUSTRALIA, TELEPHONE (08) 8259 5555